# Constacyclic codes as invariant subspaces

Radkova, D.; van Zanten, A.J.

*Published in:*
Linear Algebra and its Applications

*Publication date:*
2009

*Document Version*
Publisher's PDF, also known as Version of record

Linear Algebra and its Applications 430 (2009) 855–864

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

# Constacyclic codes as invariant subspaces

D. Radkova *, A.J. Van Zanten

*Delft University of Technology, Faculty of Information Technology and Systems, Department of Mathematics, P.O. Box 5031, 2600 GA Delft, The Netherlands*

ARTICLE INFO

*Article history:*
Received 26 May 2008
Accepted 29 September 2008

Submitted by R.A. Brualdi

*AMS classification:*
Main
94B15
Secondary
47A15

*Keywords:*
Cyclic codes
Constacyclic codes
Invariant subspaces

ABSTRACT

Constacyclic codes are generalizations of the familiar linear cyclic codes. In this paper constacyclic codes over a finite field $F$ are regarded as invariant subspaces of $F^n$ with respect to a suitable linear operator. By applying standard techniques from linear algebra one can derive properties of these codes which generalize several well-known results for cyclic codes, such as the various lower bounds for the minimum distance.



## 1. Introduction

Constacyclic codes were introduced in [2] as generalizations of linear cyclic codes. A $q$-ary constacyclic code of length $n$ can be defined by an $n \times n$-generator matrix with the property that each row (apart from the last one) $(c_0, c_1, \ldots, c_{n-1}), c_i \in \mathrm{GF}(q)$, defines the next row as $(ac_{n-1}, c_1, \ldots, c_{n-2})$, where $a$ is some fixed element from $\mathrm{GF}(q) \setminus \{0\}$. Special subclasses are the cyclic codes ($a = 1$) and the negacyclic codes ($a = -1$). In [3] an alternative point of view is taken by regarding constacyclic codes as a certain kind of contractions of cyclic codes.

Cyclic codes are traditionally described by using methods of commutative algebra (cf. e.g. [1, Chapter 7]). In this approach a codeword $(c_0, c_1, \ldots, c_{n-1})$ corresponds to a polynomial $c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$

* Corresponding author.
*E-mail address:* dradkova@fmi.uni-sofia.bf (D. Radkova).

0024-3795/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2008.09.036

which is in $R_n[x]$, the ring of polynomials in $x \bmod x^n - 1$. A cyclic shift of a codeword then corresponds to multiplication of the polynomial by $x$, and hence the theory of linear cyclic codes comes down to studying principal ideals in $R_n[x]$ generated by some generator polynomial.

This standard approach of cyclic codes seems not very appropriate for generalization to constacyclic codes in general. Since linear codes have the structure of linear subspaces of $\mathrm{GF}(q)^n$, an alternative description of constacyclic codes in terms of linear algebra appears to be another quite natural setting. In this paper we develop such an approach. Our starting point will be the characteristic polynomial of the matrix which represents the constacyclic transformation with respect to $a$ in the linear space $\mathrm{GF}(q)^n$. Another major tool is an application of the theorem of Cayley–Hamilton. This approach enables us to derive some properties for the corresponding idempotent matrices of constacyclic codes and to obtain lower bounds for the minimum distance of constacyclic codes that are generalizations of the well-known BCH, Hartmann–Tzeng and Roos bounds for cyclic codes (cf. [1]).

Throughout this paper we require that $(n, q) = 1$, which is common practice in the theory of cyclic codes.

## 2. Linear constacyclic codes as invariant subspaces

Let $F = \mathrm{GF}(q)$ and let $F^n$ be the $n$-dimensional vector space over $F$ with the standard basis $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 1)$.

Let $a$ be a nonzero element of $F$ and let

$$\psi_a : \begin{cases} F^n \to F^n \\ (x_1, x_2, \ldots, x_n) \mapsto (ax_n, x_1, \ldots, x_{n-1}). \end{cases} \tag{2.1}$$

Then $\psi_a \in \mathrm{Hom}F^n$ and it has the following matrix:

$$A(n, a) = A = \begin{pmatrix} 0 & 0 & 0 & \cdots & a \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{2.2}$$

with respect to the basis $e = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$. Note that the relations $A^{-1} = A^t$ and $A^n = aE$ hold. The characteristic polynomial of $A$ is

$$f_A(x) = \begin{vmatrix} -x & 0 & 0 & \cdots & a \\ 1 & -x & 0 & \cdots & 0 \\ 0 & 1 & -x & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -x \end{vmatrix} = (-1)^n(x^n - a). \tag{2.3}$$

In the next we shall denote (2.3) by $f(x)$. For our purposes we need the following well-known fact.

**Proposition 1.** *Let $\varphi \in \mathrm{Hom}V$ and let $U$ be a $\varphi$-invariant subspace of $V$ and $\dim_F V = n$. Then $f_{\varphi|_U}(x)$ divides $f_\varphi(x)$. In particular, if $V = U \oplus W$ and $W$ is a $\varphi$-invariant subspace of $F^n$ then $f_\varphi(x) = f_{\varphi|_U}(x) f_{\varphi|_W}(x)$.*

Let $f(x) = (-1)^n f_1(x) \cdots f_t(x)$ be the factorization of $f(x)$ into irreducible factors over $F$. According to the Theorem of Cayley–Hamilton the matrix $A$ of (2.2) satisfies

$$f(A) = O. \tag{2.4}$$

We assume that $(n, q) = 1$. In that case $f(x)$ has distinct factors $f_i(x)$, $i = 1, \ldots, t$, which are monic. Furthermore, we consider the homogeneous set of equations

$$f_i(A)\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in F^n \tag{2.5}$$

for $i = 1, \ldots, t$. If $U_i$ stands for the solution space of (2.5), then we may write $U_i = \mathrm{Ker} f_i(\psi_a)$.

**Theorem 1.** *The subspaces $U_i$ of $F^n$ satisfy the following conditions*:

(1) *$U_i$ is a $\psi_a$-invariant subspace of $F^n$*;
(2) *if $W$ is a $\psi_a$-invariant subspace of $F^n$ and $W_i = W \cap U_i$ for $i = 1, \ldots, t$, then $W_i$ is $\psi_a$-invariant and $W = W_1 \oplus \cdots \oplus W_t$*;
(3) $F^n = U_1 \oplus \cdots \oplus U_t$;
(4) $\dim_F U_i = \deg f_i(x) = k_i$;
(5) $f_{\psi_a|_{U_i}}(x) = (-1)^{k_i} f_i(x)$;
(6) *$U_i$ is a minimal $\psi_a$-invariant subspace of $F^n$*.

The proofs for the various statements of Theorem 1 are elementary and straightforward. For the details we refer to [6].

**Proposition 2.** *Let $U$ be a $\psi_a$-invariant subspace of $F^n$. Then $U$ is a direct sum of some of the minimal $\psi_a$-invariant subspaces $U_i$ of $F^n$.*

**Proof.** This follows immediately from property (2) of Theorem 1. □

**Definition 1.** A linear code of length $n$ and rank $k$ is a linear subspace $C$ with dimension $k$ of the vector space $F^n$.

**Definition 2.** Let $a$ be a nonzero element of $F$. A code $C$ with length $n$ over $F$ is called constacyclic with respect to $a$, if whenever $\mathbf{x} = (c_1, c_2, \ldots, c_n)$ is in $C$, then so is $\mathbf{y} = (ac_n, c_1, \ldots, c_{n-1})$.

The following statement will be clear from the definition.

**Proposition 3.** *A linear code $C$ of length $n$ over $F$ is constacyclic iff $C$ is a $\psi_a$-invariant subspace of $F^n$.*

**Theorem 2.** *Let $C$ be a linear constacyclic code of length $n$ over $F$. Then the following facts hold.*

(1) *$C = U_{i_1} \oplus \cdots \oplus U_{i_s}$ for some minimal $\psi_a$-invariant subspaces $U_{i_r}$ of $F^n$ and $k := \dim_F C = k_{i_1} + \cdots + k_{i_s}$, where $k_{i_r}$ is the dimension of $U_{i_r}$*;
(2) $f_{\psi_a|_C}(x) = (-1)^k f_{i_1}(x) \cdots f_{i_s}(x) = g(x)$;
(3) $\mathbf{c} \in C$ *iff* $g(A)\mathbf{c} = \mathbf{0}$;
(4) *the polynomial $g(x)$ has the smallest degree with respect to property* (3);
(5) $\mathrm{rank}(g(A)) = n - k$.

**Proof.** (1) This follows from Proposition 2.

(2) Let $(\mathbf{g}_1^{(i_r)}, \ldots, \mathbf{g}_{k_{i_r}}^{(i_r)})$ be a basis of $U_{i_r}$ over $F$, $r = 1, \ldots, s$, and let $A_{i_r}$ be the matrix of $\psi_a|_{U_{i_r}}$ with respect to that basis. Let $\tilde{f}_i(x) = f_{\psi_a|_{U_{i_r}}}(x)$. Then $(\mathbf{g}_1^{(i_1)}, \ldots, \mathbf{g}_{k_{i_1}}^{(i_1)}, \ldots, \mathbf{g}_1^{(i_s)}, \ldots, \mathbf{g}_{k_{i_s}}^{(i_s)})$ is a basis of $C$ over $F$ and $\psi_a|_C$ is represented by the following matrix:

$$\begin{pmatrix} A_{i_1} & & & \\ & A_{i_2} & & \\ & & \ddots & \\ & & & A_{i_s} \end{pmatrix}$$

with respect to that basis. Hence,

$$f_{\psi_a|_C}(x) = \tilde{f}_{i_1}(x) \cdots \tilde{f}_{i_s}(x) = (-1)^{k_{i_1} + \cdots + k_{i_s}} f_{i_1}(x) \cdots f_{i_s}(x).$$

(3) Let $\mathbf{c} \in C$. Then $\mathbf{c} = \mathbf{u}_{i_1} + \cdots + \mathbf{u}_{i_s}$ for some $\mathbf{u}_{i_r} \in U_{i_r}$, $r = 1, \ldots, s$, and $g(A)\mathbf{c} = (-1)^k[(f_{i_1} \cdots f_{i_s})(A)\mathbf{u}_{i_1} + \cdots + (f_{i_1} \cdots f_{i_s})(A)\mathbf{u}_{i_s}] = \mathbf{0}$.

Conversely, suppose that $g(A)\mathbf{c} = \mathbf{0}$ for some $\mathbf{c} \in F^n$. According to Theorem 1 we have that $\mathbf{c} = \mathbf{u}_1 + \cdots + \mathbf{u}_t$, $\mathbf{u}_i \in U_i$. Then $g(A)\mathbf{c} = (-1)^k[(f_{i_1} \cdots f_{i_s})(A)\mathbf{u}_1 + \cdots + (f_{i_1} \cdots f_{i_s})(A)\mathbf{u}_t] = \mathbf{0}$, so that $g(A)(\mathbf{u}_{j_1} + \cdots + \mathbf{u}_{j_l}) = \mathbf{0}$, where $\{j_1, \ldots, j_l\} = \{1, \ldots, t\} \backslash \{i_1, \ldots, i_s\}$. Let $\mathbf{v} = \mathbf{u}_{j_1} + \cdots + \mathbf{u}_{j_l}$ and

$$h(x) = \frac{(-1)^n(x^n - a)}{g(x)} = \frac{f(x)}{g(x)}.$$

Since $(h(x), g(x)) = 1$, there are polynomials $a(x)$, $b(x) \in F[x]$ such that $a(x)h(x) + b(x)g(x) = 1$. Hence $\mathbf{v} = a(A)h(A)\mathbf{v} + b(A)g(A)\mathbf{v} = \mathbf{0}$ and so $\mathbf{c} \in C$.

(4) Suppose that $b(x) \in F[x]$ is a nonzero polynomial of smallest degree such that $b(A)\mathbf{c} = \mathbf{0}$ for all $\mathbf{c} \in C$. By the division algorithm in $F[x]$ there are polynomials $q(x), r(x)$ such that $g(x) = b(x)q(x) + r(x)$, where $\deg r(x) < \deg b(x)$. Then for each vector $\mathbf{c} \in C$ we have $g(A)\mathbf{c} = q(A)b(A)\mathbf{c} + r(A)\mathbf{c}$ and hence, $r(A)\mathbf{c} = \mathbf{0}$. But this contradicts the choice of $b(x)$ unless $r(x)$ is identically zero. Thus, $b(x)$ divides $g(x)$. If $\deg b(x) < \deg g(x)$, then $b(x)$ is a product of some of the irreducible factors of $g(x)$, and without loss of generality we may assume that $b(x) = (-1)^{k_{i_1}+\cdots+k_{i_m}} f_{i_1} \cdots f_{i_m}$ and $m < s$. Let us consider the code $C' = U_{i_1} \oplus \cdots \oplus U_{i_m} \subset C$. Then $b(x) = f_{\psi_a|_{C'}}(x)$ and by the equation $g(A)\mathbf{c} = \mathbf{0}$ for all $\mathbf{c} \in C$ we obtain that $C \subseteq C'$. This contradiction proves the statement.

(5) By property (3) $C$ is the solution space of the homogeneous set of equations $g(A)\mathbf{x} = \mathbf{0}$. Then $\dim_F C = k = n - \operatorname{rank}(g(A))$, which proves the statement. □

**Definition 3.** Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1 \ldots, y_n)$ be two vectors in $F^n$. We define an inner product over $F$ by $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + \cdots + x_ny_n$. If $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, we say that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal to each other.

**Definition 4.** Let $C$ be a linear code of length $n$ over $F$. We define the dual of $C$ (which is denoted by $C^\perp$) to be the set of all vectors which are orthogonal to all codewords in $C$, i.e.,

$$C^\perp = \{\mathbf{v} \in F^n | \langle \mathbf{v}, \mathbf{c} \rangle = 0 \; \forall \, \mathbf{c} \in C\}.$$

It is well known that if $C$ is $k$-dimensional, then $C^\perp$ is an $(n - k)$-dimensional subspace of $F^n$, so $C^\perp$ is a linear code again.

**Proposition 4.** *The dual of a linear constacyclic code with respect to $a$ is a constacyclic code with respect to $a^{-1}$.*

**Proof.** The proof follows from the equality

$$\begin{aligned}\langle \psi_a(\mathbf{c}), \mathbf{h} \rangle &= \langle A(n, a)\mathbf{c}, \mathbf{h} \rangle = \langle \mathbf{c}, A(n, a)^t \mathbf{h} \rangle \\ &= \left\langle \mathbf{c}, A\left(n, \frac{1}{a}\right)^{-1} \mathbf{h} \right\rangle = a \left\langle \mathbf{c}, \psi_{\frac{1}{a}}^{n-1}(\mathbf{h}) \right\rangle = 0\end{aligned}$$

for every $\mathbf{c} \in C$ and $\mathbf{h} \in C^\perp$. □

**Proposition 5.** *The matrix $H$ the rows of which constitute an arbitrary set of $n - k$ linearly independent rows of $g(A)$, is a parity check matrix of $C$.*

**Proof.** The proof follows from the equation $g(A)\mathbf{c} = \mathbf{0}$ for every vector $\mathbf{c} \in C$ and from the fact that $\operatorname{rank}(g(A)) = n - k$. □

## 3. Idempotent matrices for linear constacyclic codes

Let $C$ be a linear constacyclic code of length $n$ over $F$. Then $g(x) = f_{\psi_a|_C}(x)$ (cf. Theorem 2) and $h(x) = \frac{f(x)}{g(x)}$. Since $(g(x), h(x)) = 1$, by the Euclidean algorithm there are unique polynomials $u(x)$, $v(x) \in$

$F[x]$, such that

$$u(x)g(x) + v(x)h(x) = 1, \quad \deg u(x) < \deg h(x), \quad \deg v(x) < \deg g(x). \tag{3.1}$$

It follows that

$$v(x)h(x)[u(x)g(x) + v(x)h(x)] = v(x)h(x) \tag{3.2}$$

and hence

$$v(A)h(A)[u(A)g(A) + v(A)h(A)] = v(A)h(A).$$

We next introduce the polynomial $e(x) = v(x)h(x)$ and the corresponding matrix

$$e(A) = v(A)h(A). \tag{3.3}$$

Because of $h(A)g(A) = f(A) = O$ (Cayley–Hamilton) it follows that

$$e^2(A) = e(A). \tag{3.4}$$

Now let $C = U_i$. Then $g(x) = (-1)^{k_i} f_i(x)$ and $h(x) = (-1)^{n-k_i}\hat{f}_i(x)$, where $k_i = \dim_F U_i$. Let us denote $e_i(A) = (-1)^{n-k_i} v_i(A)\hat{f}_i(A)$, $i = 1, \ldots, t$.

**Theorem 3.** *The matrices $e_i(A)$, $i = 1, \ldots, t$, satisfy the following relations*:

(1) $e_i^2(A) = e_i(A)$;
(2) $e_i(A)e_j(A) = O$ *for* $j \neq i$;
(3) $\mathbf{c} \in U_i$ *iff* $e_i(A)\mathbf{c} = \mathbf{c}$;
(4) $e_i(A)\mathbf{c} = \mathbf{0}$ *for all* $\mathbf{c} \in U_j$, $j \neq i$;
(5) $\sum_{i=1}^{t} e_i(A) = E$;
(6) *the columns of* $e_i(A)$ *generate* $U_i$.

**Proof.** (1) It follows immediately from the definition of the matrices $e_i(A)$.

(2) $e_i(A)e_j(A) = (-1)^{2n-(k_i+k_j)} v_i(A)v_j(A)\hat{f}_i(A)\hat{f}_j(A) = u(A)f(A) = O$ for a suitable polynomial $u(x) \in F[x]$.

(3) Let $\mathbf{c} \in U_i$. Then from the equality $(-1)^{k_i} u_i(x)f_i(x) + (-1)^{n-k_i} v_i(x)\hat{f}_i(x) = 1$ it follows that $(-1)^{k_i} u_i(A)f_i(A)\mathbf{c} + (-1)^{n-k_i} v_i(A)\hat{f}_i(A)\mathbf{c} = e_i(A)\mathbf{c} = \mathbf{c}$. Conversely, suppose that $e_i(A)\mathbf{c} = \mathbf{c}$ for some $\mathbf{c} \in F^n$. Then

$$f_i(A)\mathbf{c} = f_i(A)e_i(A)\mathbf{c} = (-1)^{n-k_i} v_i(A)f(A)\mathbf{c} = \mathbf{0},$$

so that $\mathbf{c} \in U_i$. Here, we applied again the theorem of Cayley-Hamilton, i.e., $f(A) = O$.

(4) Let $\mathbf{c} \in U_j$, $j \neq i$. Then

$$e_i(A)\mathbf{c} = (-1)^{n-k_i} v_i(A)\hat{f}_i(A)\mathbf{c} = u(A)f_j(A)\mathbf{c} = \mathbf{0}$$

for a suitable polynomial $u(x) \in F[x]$.

(5) Let $\mathbf{u} \in F^n$, then $\mathbf{u} = \mathbf{u}_1 + \cdots + \mathbf{u}_t$, where $\mathbf{u}_i \in U_i$, $i = 1, \ldots, t$. Then according to properties (3) and (4) we have that

$$\sum_{i=1}^{t} e_i(A)\mathbf{u} = \sum_{i=1}^{t} e_i(A)\mathbf{u}_1 + \cdots + \sum_{i=1}^{t} e_i(A)\mathbf{u}_t = \mathbf{u}_1 + \cdots + \mathbf{u}_t = \mathbf{u}.$$

Hence, $\sum_{i=1}^{t} e_i(A)\mathbf{u} = \mathbf{u}$ for all $\mathbf{u} \in F^n$, so

$$\sum_{i=1}^{t} e_i(A) = E.$$

(6) Since $f_i(A)e_i(A) = O$, the columns of $e_i(A)$ are vectors in $U_i$. From the equality $e_i(A)\mathbf{c} = \mathbf{c}$ for all $\mathbf{c} \in U_i$ it follows that

$$\begin{aligned}
&e_{11}^{(i)}c_1 + e_{12}^{(i)}c_2 + \cdots e_{1n}^{(i)}c_n = c_1,\\
&e_{21}^{(i)}c_1 + e_{22}^{(i)}c_2 + \cdots e_{2n}^{(i)}c_n = c_2,\\
&\vdots\\
&e_{n1}^{(i)}c_1 + e_{n2}^{(i)}c_2 + \cdots e_{nn}^{(i)}c_n = c_n,
\end{aligned}$$

where $e_i(A) = (e_{kl}^{(i)})$ and $\mathbf{c} = (c_1, \ldots, c_n)$. If we denote by $\mathbf{E}_i$ the $i$th vector-column of $e_i(A)$, the last equalities give us that $c_1\mathbf{E}_1 + \cdots + c_n\mathbf{E}_n = \mathbf{c}$, i.e., every vector $\mathbf{c} \in U_i$ is a linear combination of the columns of $e_i(A)$. Therefore the columns of $e_i(A)$ generate $U_i$. □

**Definition 5.** The idempotent matrices from the previous theorem will be called primitive idempotent matrices.

**Theorem 4.** *The primitive idempotent matrix $e_i(A)$, $i = 1, \ldots, t$, is the only idempotent matrix satisfying $e_i(A)\mathbf{c} = \mathbf{c}$ for all $\mathbf{c} \in U_i$ and $e_i(A)\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \sum_{j\neq i} U_j$.*

**Proof.** Let $\mathscr{E}$ be some matrix with $\mathscr{E}^2 = \mathscr{E}$ and $\mathbf{c} \in U_i$ iff $\mathscr{E}\mathbf{c} = \mathbf{c}$. It follows that $\mathrm{Im}\mathscr{E} = U_i$. For each $\mathbf{x} \in F^n$ we can write

$$\mathbf{x} = \mathscr{E}\mathbf{x} + \mathbf{x} - \mathscr{E}\mathbf{x}.$$

Now $\mathscr{E}\mathbf{x} \in \mathrm{Im}\mathscr{E}$ and $\mathbf{x} - \mathscr{E}\mathbf{x} \in \mathrm{Ker}\mathscr{E}$, since $\mathscr{E}(\mathbf{x} - \mathscr{E}\mathbf{x}) = \mathscr{E}\mathbf{x} - \mathscr{E}^2\mathbf{x} = \mathbf{0}$. It is also obvious that $F^n = \mathrm{Im}\mathscr{E} \oplus \mathrm{Ker}\mathscr{E}$, and hence it follows that $\mathrm{Ker}\mathscr{E} = \sum_{j\neq i} U_j$. So, for all $\mathbf{x} \in F^n$ we have $\mathscr{E}\mathbf{x} = e_i(A)\mathbf{x}$, or equivalently $\mathscr{E} = e_i(A)$ is the matrix projecting $F^n$ on $U_i$. □

**Remark.** $e_i(A)$ is not a unique idempotent matrix satisfying the only if-part of property (3). Indeed, let us consider the matrix $e_i(A) + e_j(A)$, $j \neq i$. Then

$$(e_i(A) + e_j(A))^2 = e_i^2(A) + e_j^2(A) = e_i(A) + e_j(A)$$

and for all vectors $\mathbf{c} \in U_i$ we have

$$(e_i(A) + e_j(A))\mathbf{c} = e_i(A)\mathbf{c} + e_j(A)\mathbf{c} = \mathbf{c} + \mathbf{0} = \mathbf{c}.$$

Now let $C = U_{i_1} \oplus \cdots \oplus U_{i_s}$ be an arbitrary linear constacyclic code of length $n$ over $F$. Then $f_{\psi_a|_C}(x) = (-1)^k f_{i_1}(x) \cdots f_{i_s}(x) = g(x)$ and

$$h(x) = \frac{f(x)}{g(x)} = (-1)^{n-k} f_{j_1}(x) \cdots f_{j_l}(x), \tag{3.5}$$

where $\{j_1, \ldots, j_l\} = \{1, \ldots, t\} \backslash \{i_1, \ldots, i_s\}$.

**Theorem 5.** *Let $C = U_{i_1} \oplus \cdots \oplus U_{i_s}$ be a linear constacyclic code of length $n$ over $F$. Then the following facts hold:*

(1) $\mathbf{c} \in C$ *iff* $e(A)\mathbf{c} = \mathbf{c}$;
(2) *the columns of* $e(A)$ *generate* $C$;
(3) $e(A) = e_{i_1}(A) + \cdots + e_{i_s}(A)$;
(4) *the constacyclic code* $C' = U_{j_1} \oplus \cdots \oplus U_{j_l}$ *has the idempotent matrix* $E - e(A)$.

**Proof.** (1) Let $\mathbf{c} \in C$. Then from the equality $u(x)g(x) + v(x)h(x) = 1$ it follows that $u(A)g(A)\mathbf{c} + v(A)h(A)\mathbf{c} = e(A)\mathbf{c} = \mathbf{c}$. Conversely, suppose that $e(A)\mathbf{c} = \mathbf{c}$ for some $\mathbf{c} \in F^n$. Then $g(A)\mathbf{c} = g(A)e(A)\mathbf{c} = v(A)f(A)\mathbf{c} = \mathbf{0}$, so $\mathbf{c} \in C$.

(2) The proof is analogous to the proof of property (6) of Theorem 3.

(3) Let us denote by $E(A)$ the idempotent matrix $e_{i_1}(A)+\cdots+e_{i_s}(A)$. Since $e(A)$ and $E(A)$ are polynomials in $A$, the equality $e(A)E(A)=E(A)e(A)$ holds. If $\mathbf{c}\in C$, then $\mathbf{c}=\mathbf{u}_{i_1}+\cdots+\mathbf{u}_{i_s}$, where $\mathbf{u}_{i_r}\in U_{i_r}$, $r=1,\dots,s$, and so

$$E(A)\mathbf{c}=[e_{i_1}(A)+\cdots+e_{i_s}(A)](\mathbf{u}_{i_1}+\cdots+\mathbf{u}_{i_s})=\mathbf{u}_{i_1}+\cdots+\mathbf{u}_{i_s}=\mathbf{c},$$

according to Theorem 3. Therefore, the columns of $E(A)$ are in $C$ and $e(A)E(A)=E(A)$. On the other hand, the columns of $e(A)$ generate $C$, so $E(A)e(A)=e(A)$. Finally, we conclude that

$$e(A)=E(A)e(A)=e(A)E(A)=E(A).$$

(4) Let $C'=U_{j_1}\oplus\cdots\oplus U_{j_l}$, then $f_{\psi_a|_{C'}}(x)=(-1)^{n-k}f_{j_1}(x)\dots f_{j_l}(x)=h(x)$, which satisfies (3.5). Then according to Theorem 3 and the previous property we have that the idempotent of $C'$ is

$$e'(A)=e_{j_1}(A)+\cdots+e_{j_l}(A)=E-\sum_{r=1}^{s}e_{i_r}(A)=E-e(A)(=u(A)g(A)),$$

which proves the statement. □

## 4. Bounds for constacyclic codes

Let $K=GF(q^m)$ be the splitting field of the polynomial $f(x)=(-1)^n(x^n-a)$ over $F=GF(q)$, where $0\neq a\in F$. Let the eigenvalues of $\psi_a$ be $\alpha_1,\dots,\alpha_n$, with $\alpha_i=\sqrt[n]{a}\alpha^i$, $i=1,\dots,n$, where $\alpha$ is a primitive $n$th root of unity and $\sqrt[n]{a}$ is a fixed, but otherwise arbitrary zero of the polynomial $x^n-a$. Let $\mathbf{v}_i$ be the respective eigenvectors, $i=1,\dots,n$. More in particular we have

$$A\mathbf{v}_i^t=\alpha_i\mathbf{v}_i^t,\quad \mathbf{v}_i=(\alpha_i^{n-1},\alpha_i^{n-2},\dots,\alpha_i,1),\quad i=1,\dots,n, \tag{4.1}$$

where $A$ is the matrix of (2.2).

Let us consider the basis $v=(\mathbf{v}_1,\dots,\mathbf{v}_n)$ of eigenvectors of $\psi_a$. With respect to this basis we have $\mathbf{c}\in C$ iff $g(A)\mathbf{c}=\mathbf{0}$. We carry out the basis transformation $e\to v$, and obtain

$$D=\begin{pmatrix}\alpha_1 & 0 & \dots & 0\\ 0 & \alpha_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \alpha_n\end{pmatrix}=T^{-1}AT, \tag{4.2}$$

with

$$T=\begin{pmatrix}\alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1}\\ \alpha_1^{n-2} & \alpha_2^{n-2} & \cdots & \alpha_n^{n-2}\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_1 & \alpha_2 & \cdots & \alpha_n\\ 1 & 1 & \cdots & 1\end{pmatrix}. \tag{4.3}$$

The columns of $T$ are the transposed of the eigenvectors $\mathbf{v}_i=(\alpha_i^{n-1},\dots,\alpha_i,1)$, $i=1,\dots,n$.

Let $\mathbf{u}_i=(\alpha_i,\alpha_i^2,\dots,\alpha_i^{n-1},\alpha_i^n)$, $i=1,\dots,n$. Then

$$\langle\mathbf{v}_i,\mathbf{u}_j\rangle=a\sum_{k=1}^{n}\left(\frac{\alpha_i}{\alpha_j}\right)^k=a\sum_{k=1}^{n}(\alpha^{i-j})^k=\begin{cases}an \text{ with } i=j,\\ 0 \text{ otherwise.}\end{cases}$$

From this it follows immediately that

$$T^{-1}=\frac{1}{an}\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_2\\ \vdots\\ \mathbf{u}_n\end{pmatrix}=\frac{1}{an}\begin{pmatrix}\alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} & \alpha_1^n\\ \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} & \alpha_2^n\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} & \alpha_n{}^n\end{pmatrix}. \tag{4.4}$$

Since $D$ is a diagonal matrix, the matrices $g(D)$ and $h(D)$ are also diagonal:

$$g(D) = \begin{pmatrix} g(\alpha_1) & 0 & \cdots & 0 \\ 0 & g(\alpha_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g(\alpha_n) \end{pmatrix}, \quad h(D) = \begin{pmatrix} h(\alpha_1) & 0 & \dots & 0 \\ 0 & h(\alpha_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h(\alpha_n) \end{pmatrix}. \tag{4.5}$$

Let $\deg h(x) = n - k = r$, and let its $r$ zeros be $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$ and its $k$ nonzeros $\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_k}$. It is obvious that the zeros of $g(x)$ are the nonzeros of $h(x)$ and vice versa.

Assume that $\mathbf{c} = (c_1, c_2, \dots, c_n) \in F^n$ and let $\mathbf{c}' = T^{-1}\mathbf{c}$. We know $\mathbf{c} \in C$ iff $g(A)\mathbf{c} = \mathbf{0}$. The latter condition is equivalent to $g(D)\mathbf{c}' = T^{-1}g(A)TT^{-1}\mathbf{c} = T^{-1}g(A)\mathbf{c} = \mathbf{0}$, which, in its turn, is equivalent to $c'_{i_1} = c'_{i_2} = \cdots = c'_{i_r} = 0$. Hence, we get the following necessary and sufficient condition for $\mathbf{c}$ to be a codeword in $C$:

$$\mathbf{u}_{i_l}\mathbf{c} = 0, \quad l = 1, \dots, r. \tag{4.6}$$

We next shall derive a bound for the minimum distance of constacyclic codes, which is similar to the so-called Roos bound for cyclic codes in [5]. Our proof and notation are also very close to the proof and notation in [5].

Let $K$ be any finite field and $\mathscr{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ any matrix over $K$ with $n$ columns $\mathbf{a}_i$, $1 \leqslant i \leqslant n$. Let $C_{\mathscr{A}}$ denote the linear code over $K$ with $\mathscr{A}$ as parity check matrix. The minimum distance of $C_{\mathscr{A}}$ will be denoted as $d_{\mathscr{A}}$.

For any $m \times n$ matrix $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ with nonzero columns $\mathbf{x}_i \in K^m$ for $1 \leqslant i \leqslant n$, we define the matrix $\mathscr{A}(X)$ as

$$\mathscr{A}(X) := \begin{pmatrix} x_{11}\mathbf{a}_1 & x_{12}\mathbf{a}_2 & \dots & x_{1n}\mathbf{a}_n \\ x_{21}\mathbf{a}_1 & x_{22}\mathbf{a}_2 & \dots & x_{2n}\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1}\mathbf{a}_1 & x_{m2}\mathbf{a}_2 & \dots & x_{mn}\mathbf{a}_n \end{pmatrix}.$$

The following lemma describes how the parity check matrix $\mathscr{A}$ for a linear code can be extended with new rows in such a way that the minimum distance increases. A proof of this result is given by Roos (cf. [5]).

**Lemma 1.** *If $d_{\mathscr{A}} \geqslant 2$ and every $m \times (m + d_{\mathscr{A}} - 2)$ submatrix of $X$ has full rank, then $d_{\mathscr{A}(X)} \geqslant d_{\mathscr{A}} + m - 1$.*

**Definition 6.** A set $M = \{\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_l}\}$ of zeros of the polynomial $x^n - a$ in $K = \mathrm{GF}(q^m)$ will be called a consecutive set of length $l$ if a primitive $n$th root of unity $\beta$ and an exponent $i$ exist such that $M = \{\beta_i, \beta_{i+1}, \dots, \beta_{i+l-1}\}$, with $\beta_s = \sqrt[n]{a}\beta^s$, $i \leqslant s \leqslant i + l - 1$. In particular, one says that $M$ is a consecutive set of $n$th roots of unity if there is some primitive $n$th root of unity $\beta$ in $K$ such that $M$ consists of consecutive powers of $\beta$.

**Definition 7.** If $N = \{\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_t}\}$ is a set of zeros of the polynomial $x^n - a$, we denote by $U_N$ or by $U(\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_t})$ the matrix of size $t$ by $n$ over $K$ that has $(\alpha_{j_s}, \alpha_{j_s}^2, \dots, \alpha_{j_s}^n)$ as its $s$th row. If $N$ is a set of $n$th roots of unity, the similar matrix over $K$ will be denoted as $H_N$.

So, it is clear that $U_N$ is a parity check matrix for the constacyclic code $C$ over $F$ having $N$ as a set of zeros of $h(x)$. Let $C_N$ be the constacyclic code over $K$ with $U_N$ as parity check matrix, and let this code have minimum distance $d_N$. So, the minimum distance of $C$ is at least $d_N$, since $C$ is a subfield code of $C_N$ (cf. [5]).

**Theorem 6.** *If $N$ is a nonempty set of zeros of the polynomial $x^n - a$ and if $M$ is a set of $n$th roots of unity such that $|\overline{M}| \leqslant |M| + d_N - 2$ for some consecutive set $\overline{M}$ containing $M$, then $d_{MN} \geqslant d_N + |M| - 1$.*

**Proof.** Let us define $\mathscr{A} := U_N$ and $X := H_M$. Then one may easily verify that $\mathscr{A}(X) = U_{MN}$, where $MN$ is the set of all products $mn$, $m \in M$, $n \in N$. Since $N$ is nonempty, $d_{\mathscr{A}} = d_N \geqslant 2$. Hence, the assertion of the theorem follows from the lemma above if in the matrix $H_M$ every $|M| \times (|M| + d_N - 2)$ submatrix has full rank. It is sufficient to show that this is the case if $|\overline{M}| \leqslant |M| + d_N - 2$ for some consecutive set $\overline{M}$ containing $M$. Observe that $H_M$ is a submatrix of $H_{\overline{M}}$, and that in the matrix $H_{\overline{M}}$ every $|\overline{M}| \times |\overline{M}|$ submatrix is nonsingular, since the determinant of such a matrix is of Vandermonde type. So, it immediately follows that every $|M| \times |\overline{M}|$ submatrix of $H_M$ has full rank. Since $|\overline{M}| \leqslant |M| + d_N - 2$, this implies that also every $|M| \times (|M| + d_N - 2)$ submatrix of $H_M$ has full rank, which proves the theorem. □

**Corollary 1.** *Let $N$, $M$ and $\overline{M}$ be as in Theorem* 6, *with $N$ consecutive. Then $|\overline{M}| < |M| + |N|$ implies $d_{MN} \geqslant |M| + |N|$.*

**Proof.** This follows immediately from the fact that $d_N = |N| + 1$ if $N$ is a consecutive set. □

By taking for $M$ the set $\{1\}$ in Corollary 1 we obtain a generalization for constacyclic codes of the well-known BCH bound (cf. [2]).

**Corollary 2.** *Let $C$ be a linear constacyclic code of length $n$ over $F$, $g(x) = f_{\psi_a|_C}(x)$ and $h(x) = \frac{f(x)}{g(x)}$. Let for some integers $b \geqslant 1$, $\delta \geqslant 1$ the following equalities*

$$h(\alpha_b) = h(\alpha_{b+1}) = \cdots = h(\alpha_{b+\delta-2}) = 0$$

*hold, i.e., the polynomial $h(x)$ has a string of $\delta - 1$ consecutive zeros. Then the minimum distance of the code $C$ is at least $\delta$.*

If we take for $M$ also a consecutive set, Corollary 1 yields a generalization of the Hartmann–Tzeng–Roos bound (cf. [4]).

**Corollary 3.** *Let $C$ be a constacyclic code of length $n$ over $F, g(x) = f_{\psi_a|_C}(x), h(x) = \frac{f(x)}{g(x)}$, and let $\alpha$ be a primitive $n$th root of unity in $K = GF(q^m)$. Assume that there exist integers $s, b, c_1$ and $c_2$ where $s \geqslant 0$, $b \geqslant 0$, $(n, c_1) = 1$ and $(n, c_2) < \delta$, such that*

$$h(\alpha_{b+i_1c_1+i_2c_2}) = 0, \quad 0 \leqslant i_1 \leqslant \delta - 2, \quad 0 \leqslant i_2 \leqslant s.$$

*Then the minimum distance $d$ of $C$ satisfies $d \geqslant \delta + s$.*

**Example.** Let $n = 25$, $q = 7$ and $a = -1$ and let $\mu$ be a primitive 50th root of unity. Then $\mu$ is a zero of the polynomial $x^{25} + 1$. In order to classify these zeros with respect to the various irreducible polynomial divisor of $x^{25} + 1$, we first determine the cyclotomic cosets of 7 mod 50, containing the odd integers. These are

$$C_1 = \{1, 7, 49, 43\}, \quad C_3 = \{3, 21, 47, 29\}, \quad C_5 = \{5, 35, 45, 15\}, \quad C_{25} = \{25\},$$
$$C_9 = \{9, 13, 41, 37\}, \quad C_{11} = \{11, 27, 39, 23\}, \quad C_{17} = \{17, 19, 33, 31\},$$

Let the zeros of $h(x)$ be $\mu^i$ with $i \in C_1 \cup C_5 \cup C_{17}$. Since $\mu$ is a primitive 50th root of unity, it follows that $\alpha := \mu^2$ is a primitive 25th root of unity. In terms of $\alpha_i$ the zeros of $h(x)$ can be written as $\alpha_2, \alpha_3;\ \alpha_7, \alpha_8, \alpha_9;\ \alpha_{15}, \alpha_{16}, \alpha_{17};\ \alpha_{21}, \alpha_{22};\ \alpha_{24}, \alpha_{25}$. Since $h(x)$ has a string of three consecutive zeros, the linear constacyclic code $C$ defined by $h(x)$ has a minimum distance $d \geqslant 4$ according to Corollary 2. Let us consider the following two sets of three consecutive zeros: $\alpha_7, \alpha_8, \alpha_9;\ \alpha_{15}, \alpha_{16}, \alpha_{17}$. We have $c_1 = 1$, $c_2 = 8$ and $(25, 8) = 1$, and so $\delta = 4$ and $s = 1$. Therefore, Corollary 3 yields a lower bound 5 for the minimum distance $d$ of the constacyclic code $C$.

Now take $N = \{\alpha_i | i = 15, 16\}$ and $M = \{\beta^j | j = 0, 2, 3, 4\}$ with $\beta = \alpha^3$. Then the elements of $MN$ are zeros of $h(x)$. Since $d_N = 3$ and $|\overline{M}| = 5 \leqslant |M| + d_N - 2 = 4 + 3 - 2$, Theorem 6 implies that $d \geqslant d_{MN} \geqslant |M| + d_N - 1 = 4 + 3 - 1 = 6$.

## References

[1] F.G. MacWilliams, N.J.A. Sloane, The Theory of Error Correcting Codes, North-Holland, Amsterdam, The Netherlands, 1977.
[2] E.R. Berlekamp, Algebraic Coding Theory, Mc Graw-Hill Book Company, New York, 1968.
[3] J. Bierbrauer, Introduction to Coding Theory, Chapman and Hall, CRC, Boca Raton, 2005.
[4] C. Roos, A generalization of the BCH bound for cyclic codes, including the Hartmann–Tzeng bound, J. Comb. Theory Ser. A 33 (1982) 229–232.
[5] C. Roos, A new lower bound for the minimum distance of a cyclic code, IEEE Trans. Inform. Theory 29 (1983) 330–332.
[6] D. Radkova, A. Bojilov, A.J. Van Zanten, Cyclic codes and quasi-twisted codes: an algebraic approach, Report MICC 07-08, Universiteit Maastricht, 2007.